

Amendments to the Claims:

This listing of claims will replace all prior versions and listings of claims in the subject application, and please amend the claims as follows:

1. (cancelled) A system for string matching across multiple relations in a relational database management system comprising:

generating a set of strings from a set of characters,
decomposing each string into a subset of tokens,
establishing at least two relations within said strings,
establishing a similarity threshold for said relations,
sampling said at least two relations,
correlating said relations for said similarity threshold and
returning all of said tokens which meet the criteria of said similarity threshold.

2. (New) A system for string matching across multiple relations in a relational database management system, the method comprising the steps of:

- a) providing at least two relations R_1 and R_2 ,
- b) defining at least one tuple t_1 and t_2 for each relation, R_1 and R_2 ,
- c) map each tuple to a weight vector v_t
- d) normalizing each weight vector,
- e) comparing each weight vector to a predetermined similarity threshold ϕ ,
- f) returning a text-join for any relations having a similarity threshold wherein $0 \leq \phi \leq 1$.